Web data extraction Requirement document

Suggested Code Policy:

1. Copy the Project folder structure from bellow github link.

<https://github.com/hackathonScrapping/BANK.git>

1. Use Python(Mandatory) with BeautifulSoup, Scrapy and Selenium (Feel free to use any other Python libraries)
2. Store Scrapped data with provided file name and CSV format.

Assignment 1:

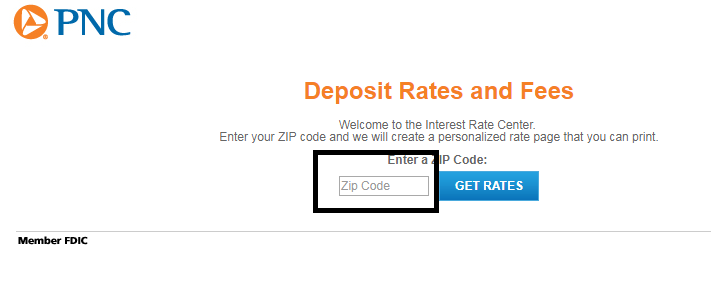
**Scarped Data Format (File Name - PNC\_Data\_Deposit\_<MM\_DD\_YYYY>.csv)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Bank Name | TAB Name | Product Name | Minimum Balance | Maximum Balance | Interest Rate | APY |
| 2/22/2018 | PNC | Spend/Reserve | Virtual Wallet | £1.00 and above | None | 0.01% | 0.01% |
| 2/22/2018 | PNC | Spend/Reserve | Virtual Wallet with Performance Spend | £2,000.00 | £9,999.99 |  |  |
| 2/22/2018 | PNC | Spend/Reserve | Virtual Wallet with Performance Spend | £10,000.00 | £49,999.99 |  |  |
| 2/22/2018 | PNC | Spend/Reserve | Virtual Wallet with Performance Spend | £50,000.00 | £99,999.99 |  |  |
| 2/22/2018 | PNC | Spend/Reserve | Virtual Wallet with Performance Spend | £100,000.00 and above | None |  |  |
| 2/22/2018 | PNC | Spend/Reserve | Virtual Wallet with Performance Spend | £1.00 and above |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **CONTINUE SIMILAR WAY FOR ALL PRODUCTCS AND BOTH TABS (Growth and Spend)** | | | | | | |  |
|  |  |  |  |  |  |  |  |
|  |  | 1 | Do not Capture Sub Product Category **Reserve Data** | |  |  |  |

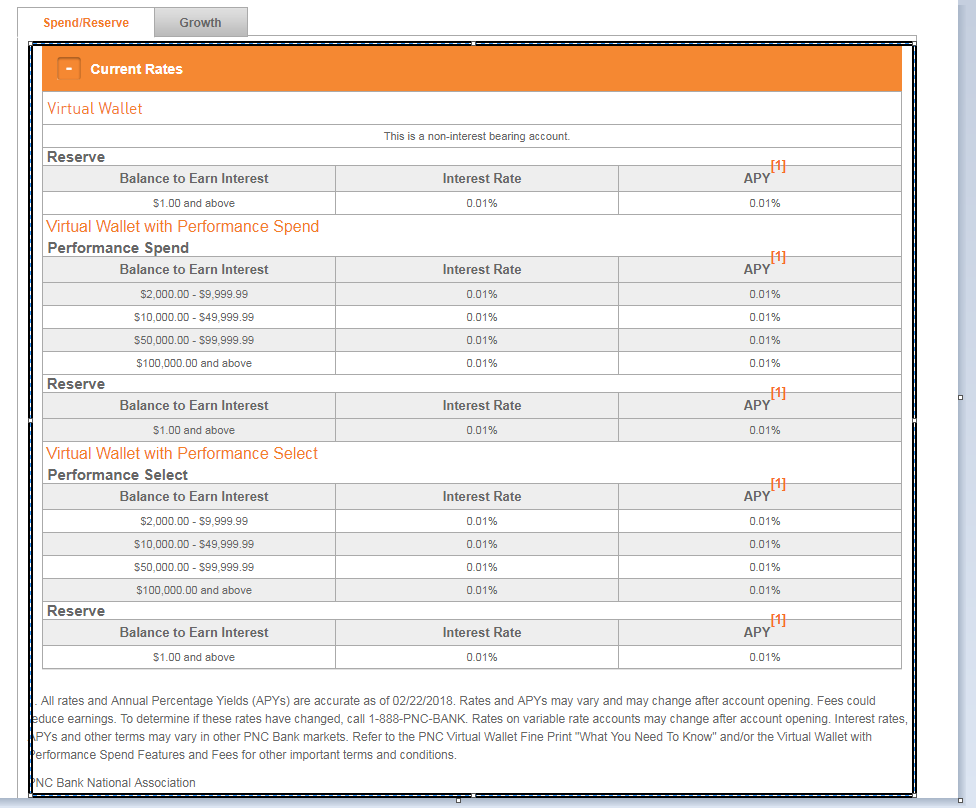
**Steps**

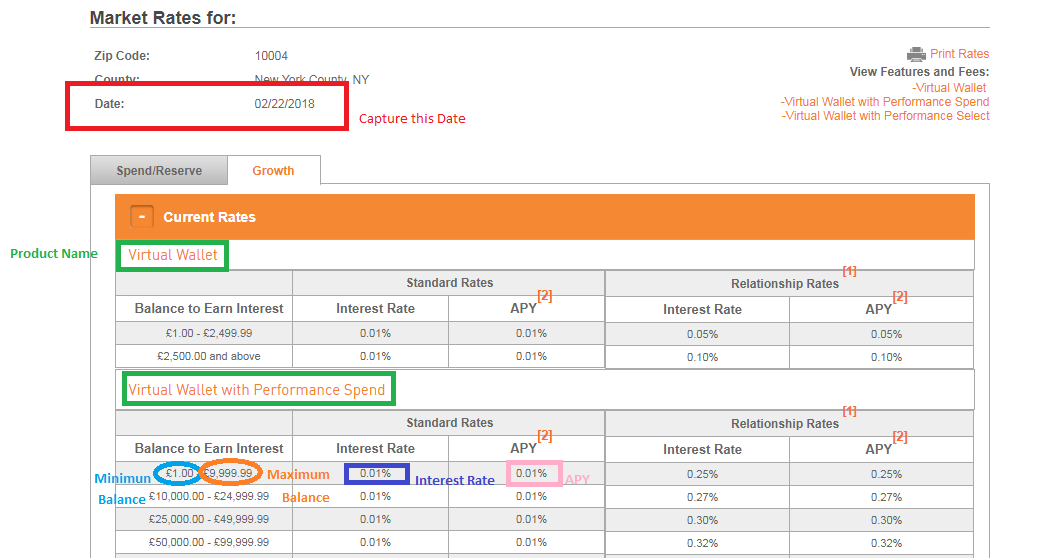
1. Click on below bank site

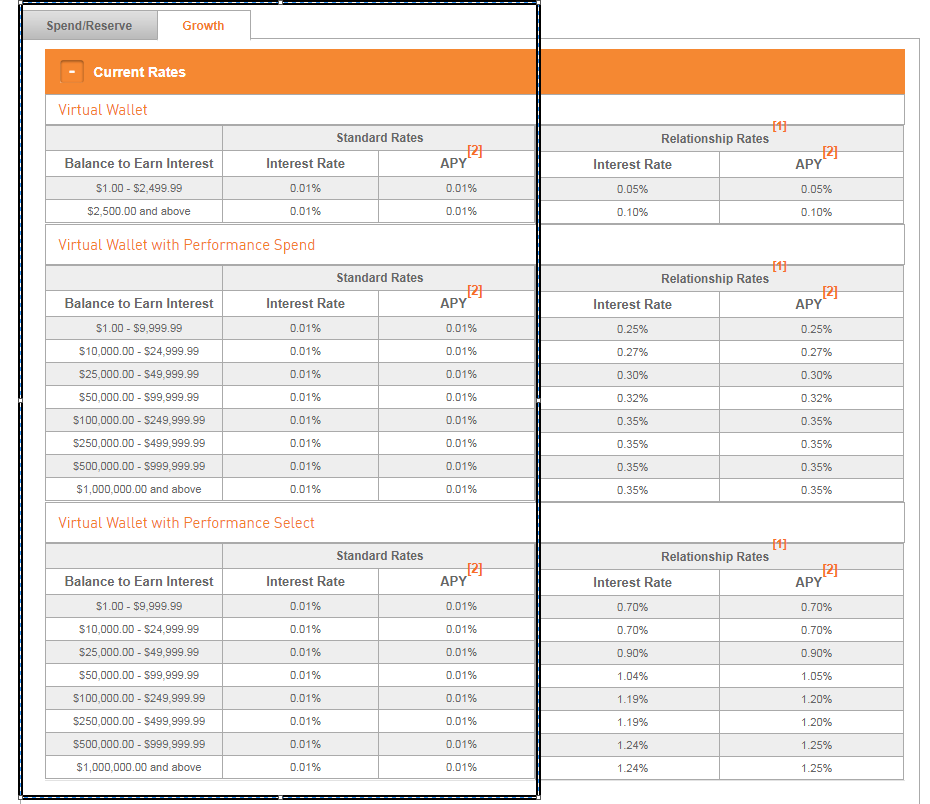
<https://apps.pnc.com/rates/servlet/DepositRatesSearchVW?productGroup=growth>



1. Input 10004 in Zipcode and click Get Rates
2. Scrape data from both “**Spend Reserve**” tab and “**Growth**” tab







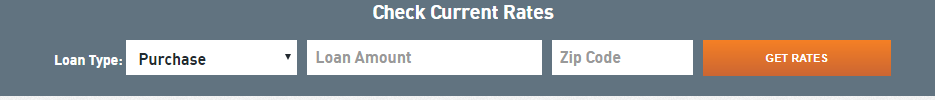
Assignment 2:

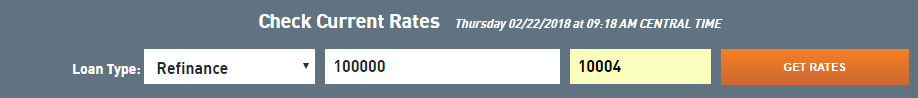
URL:

<https://www.pnc.com/en/personal-banking/borrowing/home-lending/mortgages/fixed-rate-mortgage.html>

**Scarped Data Format (File Name - PNC\_Data\_Mortgage\_<MM\_DD\_YYYY>.csv)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Date | Bank Name | Loan Type | Product Name | Interest Rate | APR |
| 2/22/2018 | PNC | Purchase | 30 Yr Fixed | 4.625% | 4.838% |
| 2/22/2018 | PNC | Purchase | 20 Yr Fixed | 4.375% | 4.838% |
| 2/22/2018 | PNC | Purchase | 15 Yr Fixed | 4.000% | 4.342% |
| 2/22/2018 | PNC | Purchase | 10 Yr Fixed | 3.875% | 4.366% |
| 2/22/2018 | PNC | Refinance |  |  |  |
|  |  |  |  |  |  |
| **Continue Same way for Refinance as well** | | | | | |





Input: Loan Amount: 100000

Zipcode: 10004

Click GET RATES

